

대한민국 특허청
KOREAN INTELLECTUAL
PROPERTY OFFICE

별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto
is a true copy from the records of the Korean Intellectual
Property Office.

출원번호 : 10-2003-0042613
Application Number

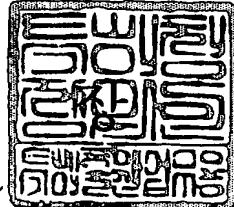
출원년월일 : 2003년 06월 27일
Date of Application JUN 27, 2003

출원인 : 주식회사 케이티
Applicant(s) KT Corporation

2004년 06월 28일

특허청

COMMISSIONER



【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0005
【제출일자】	2003.06.27
【발명의 명칭】	퍼지 기반 O C – S V M을 이용한 동영상 자동 요약 장치 및 방법
【발명의 영문명칭】	Apparatus and method for automatic video summarization using fuzzy one-class support vector machines
【출원인】	
【명칭】	주식회사 케이티
【출원인코드】	2-1998-005456-3
【대리인】	
【명칭】	유미특허법인
【대리인코드】	9-2001-100003-6
【지정된변리사】	이원일
【포괄위임등록번호】	2001-046354-9
【발명자】	
【성명의 국문표기】	최영식
【성명의 영문표기】	CHOI, YOUNG SIK
【주민등록번호】	621031-1052118
【우편번호】	412-160
【주소】	경기도 고양시 덕양구 화전동 200-1번지
【국적】	KR
【발명자】	
【성명의 국문표기】	이상윤
【성명의 영문표기】	LEE, SANG YOUN
【주민등록번호】	640522-1041512
【우편번호】	137-792
【주소】	서울특별시 서초구 우면동 17번지
【국적】	KR

【발명자】

【성명의 국문표기】 김선정
【성명의 영문표기】 KIM,SUN JEONG
【주민등록번호】 680311-2063410
【우편번호】 137-792
【주소】 서울특별시 서초구 우면동 17번지
【국적】 KR
【취지】 특허법 제42조의 규정에 의하여 위와 같이 출원합니다. 대리인
유미특허법인 (인)

【수수료】

【기본출원료】	20	면	29,000	원
【가산출원료】	7	면	7,000	원
【우선권주장료】	0	건	0	원
【심사청구료】	0	항	0	원
【합계】	36,000			원
【첨부서류】	1. 요약서·명세서(도면)_1통			

【요약서】**【요약】**

이 발명은 퍼지 기반 OC-SVM을 이용한 동영상 자동 요약 장치 및 방법에 관한 것이다.

본원 발명에서는 효과적인 비디오 요약 생성을 위해서 인간의 주관적인 판단을 반영하고, 사용자의 환경이나 요구 조건에 맞는 유연한 형태의 비디오 요약 정보를 생성하기 위한 방안을 제시하며, 주어진 비디오에서 중요한 비디오 세그먼트를 추출하고 이로부터 일련의 키 프레임들을 추출함으로써 비디오의 내용을 한눈에 파악하고 원하는 비디오 장면을 곧바로 액세스할 수 있게 한다.

【대표도】

도 2

【색인어】

비디오요약, 자동요약기법, 스키밍

【명세서】**【발명의 명칭】**

폐지 기반 OC-SVM을 이용한 동영상 자동 요약 장치 및 방법 {Apparatus and method for automatic video summarization using fuzzy one-class support vector machines}

【도면의 간단한 설명】

도1은 이 발명의 실시예에 따른 폐지 기반 OC-SVM을 이용한 동영상 자동 요약 장치의 구성도.

도2는 이 발명의 실시예에 따른 폐지 기반 OC-SVM을 이용한 동영상 자동 요약 장치의 동작 흐름도.

도3은 확장 가능(Scalable)한 요약 정보 생성에 적용하기 위한 확장 가능한 노벨티 디텍션(scalable Novelty Detection)을 설명하기 위한 개념도.

도4 및 도5는 각각 영화와 뮤직비디오에 대한 실험 결과로서, 주요 세그먼트 추출의 반복에 따른 (전체 세그먼트 대비) 증가되는 주요 세그먼트들의 비율과 이들 세그먼트들이 전체 비디오 이벤트를 커버하는 비율에 대한 도표.

【발명의 상세한 설명】**【발명의 목적】****【발명이 속하는 기술분야 및 그 분야의 종래기술】**

☞ 이 발명은 비디오 요약 기술에 관한 것으로, 특히, 폐지 기반 OC-SVM을 이용한 동영상 자동 요약 장치 및 방법에 관한 것이다.

<6> 대용량 멀티미디어 데이터베이스의 증가와 통신 및 디지털 미디어 처리 기술의 발달로 인하여 수많은 동영상 비디오가 출현함에 따라, 축약된 동영상 비디오의 요약 정보를 바탕으로 한 검색 서비스를 통해 사용자의 편의와 만족을 도모하기 위한 노력이 증대되고 있다.

<7> 그러나 현재까지 대부분의 비디오 축약은 사람이 직접 개입하여 적절한 장면이나 이미지를 추려내는 과정을 통하여 이루어져왔다.

<8> 최근 동영상 비디오와 관련된 다양한 형태의 사업이 전개되면서 대량의 동영상 비디오를 자동으로 분석하고자 하는 필요성이 증대되고 있으며, 이에 따라 이러한 문제를 해결하기 위한 많은 연구들이 활발하게 진행되고 있다.

<9> 동영상 비디오를 축약하는 방법에는 비디오 스키밍(Video skimming), 하이라이트 (highlight) 와 비디오 요약(summary) 방법이 있다.

<10> 비디오 스키밍은 비디오와 오디오 데이터로부터 추출된 중요한 의미를 갖는 부분들을 연속적으로 연결하여 짧은 비디오 시놉시스를 생성하는 방법이다.

<11> 하이라이트는 주로 특정 이벤트 중심으로 동영상의 흥미로운 일부분만 추려내어 축약한 것이다. 비디오 요약은 동영상 비디오에 대한 의미있는 내용과 구조적인 정보를 추려내기 위한 것으로, 비디오 요약의 결과는 일반적으로 주요 정지영상(키 프레임)의 연속된 형태(sequence of key frames) 표현되는데, 현재까지 대부분의 동영상 축약에 대한 연구는 주로 동영상 요약 정보 생성을 그 목표로 하고 있다.

<12> 키 프레임으로 표현된 비디오 요약은 비디오 전체의 내용을 한눈에 파악할

수 있게 할 뿐만 아니라, 그 키 프레임을 담고 있는 장면의 (scenes 또는 shot) 창구 (entries) 역할을 하게 된다. 이 때문에 비디오 요약 작업을 최적의 키 프레임 선정, 혹은 최적의 키 프레임이 위치한 세그먼트 선정 작업이라고 볼 수 있으며, 보통 색상, 움직임 등 시각적으로 도드라진 특징들이 키 프레임을 선택하는 데에 중요한 요건으로 이용된다.

<13> 한편, 비디오 요약은 그 응용범위에 따라서 샷 수준(Shot-Based)의 요약과 세그먼트 수준(Segment-Based)의 요약으로 나눌 수 있다.

<14> 샷 기반 요약은 비교적 작은 양의 비디오, 즉, 비디오 클립을 (video clip) 소수의 키 프레임으로 나타내는 방법이며, 세그먼트 기반 요약은 긴 비디오 전체를 요약하는 기술이다.

<15> 세그먼트 기반 요약 기술의 응용 범위가 더 넓은 이유로, 최근 들어서 세그먼트 기반 요약에 대한 연구가 더욱 활발히 진행되고 있다. 제안된 본 발명도 세그먼트 기반 요약을 목표로 하고 있다.

<16> 분할된 비디오를 세그먼트 단위로 요약하는 방법은 (a) 시간 윈도우 상에서 샷 간의 상호 관련성을 분석하여, 상호 관련성이 높은 샷들을 하나의 스토리 단위 (Story Unit, 혹은 Scene)로 묶는 샷 그룹핑 (Shot Grouping) 방법과 (b) 전통적인 클러스터링 방법에서 얻어진 클러스터의 성질을 분석하여 중요한 클러스터를 선택하는 방법이 있다.

<17> 이들 방법들은 어떠한 시각 특징을 사용하는 지, 혹은 어떠한 샷을 대표로 선택하는 지에 따라서 더욱 세분화될 수 있다.

<18> 그러나, 이들 방법들이 공통적으로 안고 있는 중요한 문제점이 있는데 그것은 대표성 판단 시 지나치게 임계치에 의존한다는 것이다. 즉, 특정 임계치를 두고 이에 근거하여 샷의 대표성을 판단하고 있다. 예를 들면, 샷의 중요도가 임계치 이상인 것들, 혹은 중요도가 상위

10% 순위에 있는 것들로 샷을 선정하게 된다. 더욱이, 이러한 임계치는 실험적으로 결정하고 있다. 그러나, 비디오 요약 알고리즘이 실험적 임계치에 지나치게 의존하게 됨으로써 야기되는 문제점은, 비디오 요약 시스템이 몇몇 특정 비디오에는 상당히 효과적일 수 있으나, 다양한 장르의 비디오에 적용되기 어렵다는 것이다.

<19> 또한, 이러한 문제는 다양한 종류의 비디오 정보를 다루는 응용 분야에서는 치명적인 결함이 될 수 있으며, 또한, 실험적으로 최적의 임계치를 정하는 작업 또한 상당한 비용을 초래하게 된다.

<20> 한편, 동영상 비디오 요약을 위해 키 프레임을 선택하는 데에는, 가시적인 시각적 특징 뿐만 아니라 인간의 판단에 근거한 주관적인 판단이 중요한 요소로 작용할 수 있다.

<21> 실제로 사람이 수동으로 비디오 축약을 수행하는 경우, 이러한 주관적 판단을 이입함으로써 훨씬 더 공감할 수 있는 동영상 요약을 만들어낼 수 있게 된다. 그러므로 효과적인 비디오 요약 생성을 위해서는 이러한 주관적인 판단을 동영상 축약 과정에 반영할 수 있는 연구가 필요하다.

<22> 또한, 보다 효율적으로 비디오 요약을 생성하기 위해서는 사용자의 환경을 고려하여 확장 가능한 비디오 요약 정보를 생성하는 방안도 강구되어야 한다.

【발명이 이루고자 하는 기술적 과제】

<23> 본 발명이 이루고자 하는 기술적 과제는 주어진 비디오에서 중요한 비디오 세그먼트를 추출하고 이로부터 일련의 키 프레임들을 추출함으로써 비디오의 내용을 한눈에 파악하고 원하는 비디오 장면을 곧바로 액세스할 수 있게 하는 비디오 요약 생성 기술을 제공하는 것이다.

<24> 이를 위하여 본 발명에서는 기존의 방법들에서 제기되었던 여러 문제들을 해결하기 위하여, (a) 획일화된 임계치에 의존적인 방법을 탈피하고, (b) 인간의 주관적은 판단을 반영할 수 있으며, (c) 사용자의 환경이나 요구 조건에 맞는 유연한 형태의 비디오 요약 정보를 생성하기 위한 페지 기반 One-Class SVM을 이용한 동영상 자동 요약 장치 및 방법을 제공한다.

【발명의 구성】

<25> 이러한 과제를 해결하기 위한 본 발명의 하나의 특징에 따른 페지 기반 OC-SVM을 이용한 동영상 자동 요약 장치는,

<26> 사용자로부터 요약하고자 하는 동영상 비디오 소스와 원하는 요약 시간을 입력받기 위한 입력부;

<27> 전체 비디오 소스를 샷 단위로 추출하는 샷 검출 모듈;

<28> 동영상 비디오의 장르별 특성 및 원하는 요약의 목적에 따라 중요도를 산출하는 중요도 측정 모듈;

<29> 샷 정보와 중요도 값을 페지 OC-SVM 알고리즘에 적용하여 비디오 요약 (VS)을 생성하는 비디오 요약 생성 모듈;

<30> 생성된 비디오 요약을 화면상에 출력하는 출력부;

<31> 생성된 비디오 요약을 저장하는 저장부를 포함한다.

<32> 이러한 과제를 해결하기 위한 본 발명의 하나의 특징에 따른 페지 기반 OC-SVM을 이용한 동영상 자동 요약 방법은,

<33> 사용자로부터 요약하고자 하는 동영상 비디오 소스와 원하는 요약 시간을 입력받는 제1 단계;

<34> 상기 비디오 소스를 샷 단위로 추출하는 제2 단계;

<35> 동영상 비디오의 장르별 특성 및 원하는 요약의 목적에 따라 중요도를 산출하는 제3 단계;

<36> 샷 정보와 중요도 값을 폐지 OC-SVM 알고리즘에 적용하여 비디오 요약을 생성하는 제4 단계;

<37> 생성된 비디오 요약을 화면상에 출력하는 제5 단계;

<38> 생성된 비디오 요약을 저장하는 제6 단계를 포함한다.

<39> 본 발명은 우선 획일화된 임계치에 의존적인 방법에서 벗어나 비디오 요약을 “Novelty Detection” 문제로 해석하는 개념에서 출발한다. 즉, 주어진 비디오에서 시각적 특징이 다른 프레임과 구별되는 프레임을 그 비디오의 대표되는 프레임으로 정의하는 것이다.

<40> 노벨티 디텍션(Novelty Detection) 문제에는 OC-SVM이 탁월한 성능을 보이고 있다. 원래 SVM(Support Vector Machine)은 레이블된 데이터를 이용하여 머신(machine)을 학습시킨 후 레이블되지 않은 새로운 데이터가 들어왔을 때 올바른 답을 도출해내는 확률적 학습 이론(statistical learning theory)에 기반한 학습 방법의 일종이다.

<41> 전통적인 학습 방법과 달리, 구조적 리스크를 최소화함으로써, 벡터 공간에서의 최적의 결정 경계 영역을 찾아내는 방법을 사용하여, 주로 패턴 인식과 같은 이진 분류 문제에서 최근 좋은 결과를 보여주고 있다. 이중에서도 One-class SVM(OC-SVM)은 SVM의 결정 경계 영역을 기반으로 한 퍼지티브/네거티브(positive/negative)의 두 클래스 중에서, 레이블된 퍼지티브 클래스의 데이터, 즉, 목적하는 바를 가장 잘 표현할 수 있는 데이터를 대상으로 한다. 예를 들어, 텍스트 검색이나 이미지 검색과 같은 경우, 전체 검색 대상 데이터 중에, 사용자가 관심

있어하는 대상 데이터의 양은 극히 미미하다. 이러한 경우 중요한 것은 전체 데이터가 아니라 관련된 일부 퍼지티브 샘플이다.

<42> 이러한 OC-SVM은 주어진 데이터의 분포를 예측하는 것이 아니라, 주어진 데이터를 가장 잘 설명할 수 있는 최적의 서포트 벡터(support vectors)를 예측하는 기술이다.

<43> 이렇게 OC-SVM은 일반적인 특성 벡터 공간에서 보다 특이한 특성 벡터를 찾아내는 데에 유용하므로, 동영상 비디오의 수많은 프레임 중에서 독특한 특징을 가지는 키 프레임을 효과적으로 탐지할 수 있게 된다.

<44> 하지만 OC-SVM을 통해 원하는 비디오 요약을 얻는 데에는 한계가 있다. OC-SVM은 단순한 시각적 특성을 반영한 서포트 벡터를 그 대상으로 하므로, 인간이 판단할 수 있는 주관적인 요소 - 예를 들어 짧게 처리된 것보다는 길게 처리된 장면이 더 중요한 의미를 전달할 것이라거나 뉴스의 경우 앵커맨이 들어 있는 세그먼트는 제외하고 요약 정보를 생성하는 것이 더 낫다는 등의- 를 반영하기가 어려우며, 설령 이를 반영하는 특성 벡터를 정의하더라도 이를 기준의 시각적 특징 벡터와 결합하는 과제가 남게 된다.

<45> 본 발명에서는 퍼지(fuzzy) 이론의 멤버쉽 함수 정의를 통하여 이들 사용자의 주관적 중요성을 반영하면서, 통계적으로 특징적인 벡터를 결합한 퍼지 기반 OC-SVM을 통하여 전체적인 서포트 벡터를 구한다.

<46> 한편, 퍼지 OC-SVM은 주어진 데이터를 둘러싸는 최소 구(minimum sphere)를 찾는 성질을 이용하면, 사용자의 환경을 고려한 유연한 형태의 확장 가능한 비디오 요약을 생성할 수 있다. 즉, 가장 밖에 있는 표면상의 벡터를 먼저 추출하고, 이를 최상위 요약으로 생성한다. 다음에는 이 표면을 벗겨내고, 다음 번으로 현재의 데이터를 둘러싸는 구를 찾는다. 이렇게 해서 얻

은 벡터를 먼저 구한 벡터 집합과 합침으로서 보다 상세한 비디오 요약을 생성할 수 있게 된다. 사용자의 환경에 따라 이러한 과정을 적절히 반복함으로써 확장 가능한 비디오 요약을 생성할 수 있다. 이러한 확장성은 계층적 클러스터링 알고리즘에서도 가능하다.

<47> 그러나 본 발명에서 제안된 방법으로 각 레벨에서 필요한 벡터의 수가 최적으로 구해지는 반면, 계층적 클러스터링 알고리즘에서는 클러스터링의 조건에 따라 확장 작업이 이루어지는 차이가 있다.

<48> 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였다. 명세서 전체를 통하여 유사한 부분에 대해서는 동일한 도면 부호를 붙였다.

<49> 도1은 이 발명의 실시예에 따른 퍼지 기반 OC-SVM을 이용한 동영상 자동 요약 장치의 구성도이다.

<50> 도1을 참조하면, 입력부(40), 샷검출 모듈(10), 중요도 측정 모듈(20), 비디오 요약 생성 모듈(30), 저장부(60), 출력부(50)를 포함한다. 입력부(40)는 우선 사용자로부터 두 가지 정보, 즉 ①요약하고자 하는 동영상 비디오 소스와, ② 원하는 요약 시간을 입력받는다. 샷검출 모듈(10)은 영화와 같은 긴 분량의 비디오는 주로 세그먼트 수준의 요약을 하게 되므로 이를 위한 선행 작업으로써, 전체 비디오를 샷 단위로 추출하는 작업을 샷 검출 모듈에서 수행한다. 만일 비디오 클립(video clip)과 같은 작은 양의 비디오를 주 대상으로 하는 경우에는 샷 검출 모듈을 거치지 않을 수도 있다. 샷 검출 모듈은 ①의 요약하고자 하는 동영상 비디오 소스로부터 샷을 검출하여 샷 리스트를 구성하고 이를 비디오 요약 생성 모듈(30)로 보

낸다. 중요도 측정 모듈(20)은 동영상 비디오의 장르별 특성, 혹은 원하는 요약의 목적에 따라 중요도(β)를 산출한다. 이 작업은 사람이 수동으로 데이터를 입력하거나 필요한 DB로부터 자료를 가져오는 등, 그 목적에 따라 다양한 방식으로 수행될 수 있다. 비디오 요약 생성 모듈(30)은 샷 정보와 중요도 값을 퍼지 OC-SVM 알고리즘에 적용하여 비디오 요약 (VS)을 생성한다. 이때 사용자로부터 입력으로 받은 요약 시간 정보를 바탕으로 확장성(scability) 처리를 할 것인가를 결정한다. 필요한 경우 확장성 처리 모듈(32)을 반복 수행함으로써, 사용자가 원하는 시간의 범위를 갖는 비디오 요약을 생성한다. 출력부(50)는 생성된 비디오 요약을 화면 상에 출력하고 저장부(60)는 그 정보를 저장한다.

<51> 이러한 구성을 가진 이 발명의 실시예에 따른 퍼지 기반 OC-SVM을 이용한 동영상 자동 요약 장치의 동작에 대해 상세히 설명하면 다음과 같다.

<52> 도2는 이 발명의 실시예에 따른 퍼지 기반 OC-SVM을 이용한 동영상 자동 요약 장치의 동작흐름도로서, 퍼지(Fuzzy) 기반 OC-SVM을 이용한 동영상 자동 요약 기법”을 적용하여 요약 작업을 수행하는 전 과정을 기술한 것이다.

<53> 도2를 참조하여 각 작업별로 동작을 상세히 설명한다.

<54> (0) 선행작업

<55> 비디오 요약 이전에 선행되어야 할 작업은, 일련의 비디오 프레임들을 분석하여 샷 단위로 추출하는 작업이며, 이는 샷 검출 모듈(10)에 의해 수행된다. 이는 비디오 컨텐츠를 구성하는 정형화된 개별 프레임 이미지로부터, 시각적 특성을 추출하고, 이를 바탕으로 각각의 프레임간의 유사도를 비교함으로써 시간상 혹은 공간상으로 연속적인 동작을 표현하는 샷이라는 단위로 분할하는 작업이다. 그러나, 이러한 샷 경계선 검출이 (Shot Boundary Detection) 완벽

하지는 못하기 때문에 비디오 분할을 하지 않고 프레임 단위에서부터 전통적인 HAC (Hierarchical Agglomerative Clustering)이나 K-Means Clustering 방법을 사용하기도 한다. 향후 설명할 요약 작업은 이렇게 분할된 일련의 세그먼트를 기반으로 한다.

<56> (1) 중요도 측정

<57> 중요도 검출모듈(20)은 중요도를 검출하며, 중요도는 비디오 세그먼트에 대한 사용자의 주관적 관점을 반영하여 사용자에 따라 다양하게 정의될 수 있다. 본 발명에서 정의된 중요도 (β)는 해당 세그먼트에 대한 중요도를 나타내며, 그 범위는 ($\beta \in [0, 1]$)로 한다. 이 장에서는 이해를 돋기 위해 몇 가지 예시를 통해 이를 설명한다.

<58> (1-1) 세그먼트의 길이를 고려한 경우

<59> 비디오 세그먼트의 평균길이를 m , 그리고 표준 편차를 σ 라고 할 때, 다음수학식1과 같이 세그먼트에 대한 중요도를 나타낼 수 있다.

$$\beta_j = 0.5 \left[\frac{(duration - m)}{3\sigma} + 1 \right]$$

【수학식 1】

<61> 여기서, duration은 세그먼트 j 의 길이를 나타낸다. 한편, β 가 0보다 작거나 같으면 아주 작은 값으로 (예를 들어 0.001), β 가 1보다 크면 1로 조정한다.

<62> (1-2) 사용자의 선호도에 따라 지정된 경우

<63> 예를 들어, 사용자가 비디오 세그먼트 중에서 특정 세그먼트를 지정한다거나, 아니면, 다른 비디오 소스에서 미리 저장된 비디오 세그먼트를 지정한 후, 지정된 세그먼트와 같은 세그먼트가 비디오 요약에 포함되지 않기를 원할 경우 중요도는 다음 수학식2와 같이 결정될 수 있다.

$$\beta_j = 1 - sim(x_j, u)$$

【수학식 2】

<65>

<66> 여기서, $\text{sim}(x, y)$ 는 특징 벡터 x, y 로 나타내어지는 두 세그먼트 간의 유사도이며, u 는 사용자가 지정한 세그먼트의 특징 벡터이다.

<67> (2) Fuzzy One-Class Support Vector Machine 알고리즘

<68> (2-1) 기존의 OC-SVM 알고리즘

<69> OC-SVM 알고리즘은 다음과 같다. 시각적 특성 벡터를 표현하는 데이터 포인트 중 n 개의 데이터 포인트 $\{x_j, j = 1, \dots, n\}$ 를 포함하는 데이터 집합 S 를 가정하고, 이러한 데이터 포인트를 feature space로 매핑하는 feature map Φ 를 $\Phi(x)$ 로 정의한다. 이러한 feature map space에서 “dot product”는 다음 수학식 3과 같이 정의된다.

<70> 【수학식 3】

$$\Phi(x_i) \cdot \Phi(x_j) = K(x_i, x_j).$$

<71> 여기서, $K(x_i, x_j)$ 는 여러 함수를 사용할 수 있으며 편의상 Gaussian Kernel 함수를 사용한다. 즉, $K(x_i, x_j) = \exp(-0.5\|x_i - x_j\|^2/\sigma^2)$ 이다.

<72> 이때 OC-SVM은 다음과 같은 목적함수(objective function)를 최소화하기 위한 것이다.

즉, 모든 feature vector $\Phi(x_j)$ 를 포함하는 최소 구를 찾는 것이다.

<73>

$$R^2 + C \sum \xi_j,$$

【수학식 4】 $\|\Phi(x_j) - a\|^2 \leq R^2 + \xi_j \text{ and } \xi_j \geq 0 \quad \forall j.$

<74>

Lagrangian multiplier $\alpha_j \geq 0$ and $\mu_j \geq 0$ 를 도입하면, 수학식4는 다음과 같은 L에 관한 식으로 표현된다.

<75> 【수학식 5】 $L = R^2 - \sum_j (R^2 + \xi_j - \|\phi(x_j) - a\|^2) \alpha_j - \sum_j \xi_j \mu_j + C \sum_j \xi_j$.

<76> L 을 R , a , 과 각각 미분한 결과를 0으로 놓고 풀면, 다음과 같은 dual problem을 얻을 수 있다.

<77> 【수학식 6】 Maximize

$$W = \sum_i \phi(x_i) \cdot \phi(x_i) \alpha_i - \sum_{i,j} \alpha_i \alpha_j \phi(x_i) \cdot \phi(x_j) = \sum_j \alpha_j K(x_j, x_j) - \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j)$$

<78> 여기서, $0 \leq \alpha_j \leq C$, $\sum_j \alpha_j = 1$, 및 $a = \sum_j \alpha_j \phi(x_j)$ 를 만족하여야 한다. 이때 다음의 KKT조건

<79> $\xi_j \mu_j = 0$.

【수학식 7】 $(R^2 + \xi_j - \|\phi(x_j) - a\|^2) \alpha_j = 0$.

<80> 다음과 같이 구해질 수 있다.

<81> 【수학식 8】 $R^2 = \|\phi(x) - a\|^2 = K(x, x) - 2 \sum_i \alpha_i K(x, x_i) + \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j)$.

<82> 여기서, x 는 support vector이다. 수학식 6과 같이 일반적인 Quadratic Programming으로 간단하게 a 의 값을 구할 수 있다. Quadratic Programming에서 얻어진 결과 값 a 중에서, 그 값이 0보다 크면 우리는 그에 해당하는 특징 벡터 x 를 support vector라고 부른다.

<83> (2-2) fuzzy OC-SVM 알고리즘

<84> (2-1)의 OC-SVM 알고리즘에 대해 (1)의 중요도를 결합한 fuzzy OC-SVM 알고리즘에서는 다음과 같이 최소 구를 찾을 수 있다.

<85> 다음의 데이터 집합 $S = \{(x_1, \beta_1), (x_2, \beta_2), \dots, (x_n, \beta_n)\}$ 를 가정한다. 이때 데이터 포인트 x_i 는 세그먼트에서 얻어진 시각적 특징 벡터이며, 중요도 β_i 는 해당 세그먼트에 대한 중요도를 나타낸다. 집합 S 에 퍼지 feature map을 적용한 결과는 $\{\beta_1\phi(x_1), \dots, \beta_n\phi(x_n)\}$ 가되며, 이에 따라 (2-1)의 OC-SVM에서 적용된 다양한 수식들에 대해 다음과 같이 중요도가 결합되어 유도될 수 있다.

<86> $\Phi(\beta_i, x_j) \cdot \Phi(\beta_j, x_j) = \beta_i \beta_j K(x_i, x_j)$.

【수학식 9】 $\|\beta_j\phi(x_j) - a\|^2 \leq R^2 + \xi_j$, and $\xi_j \geq 0 \quad \forall j$.

<87> Lagrangian multiplier

<88> $\alpha_j \geq 0$ and $\mu_j \geq 0$ 를 도입하여 L식으로 표현한 경우,

<89> 【수학식 10】 $L = R^2 - \sum_j (R^2 + \xi_j - \|\beta_j\phi(x_j) - a\|^2) \alpha_j - \sum_j \xi_j \mu_j + C \sum_j \xi_j$.

<90> L을 미분한 결과를 이용하여 다음과 같은 dual problem에서도,

<91> 【수학식 11】 $\text{maximize } W = \sum_j \alpha_j \beta_j^2 K(x_j, x_j) - \sum_{i,j} \alpha_i \alpha_j \beta_i \beta_j K(x_i, x_j)$
를 만족하여야 한다

<92> 또한, KKT조건은,

<93> 【수학식 12】 $(R^2 + \xi_j - \|\beta_j\phi(x_j) - a\|^2) \alpha_j = 0$.

<94> 이러한 과정을 통해 퍼지 OC-SVM을 적용하였을 때 최소 구의 반지름 R 은 다음과 같이 구해질 수 있다.

<95> 【수학식 13】

$$R^2 = \|\beta\phi(x) - a\|^2 = \beta^2 K(x, x) - 2\beta \sum_i \alpha_i \beta_i K(x, x_i) + \sum_{i,j} \alpha_i \alpha_j \beta_i \beta_j K(x_i, x_j).$$

<96> 여기서, x 는 support vector이며, β 는 해당 support vector x 의 중요도이다.

<97> (3) 비디오 요약에의 적용

<98> 폐지 OC-SVM 알고리즘 수학식 11에서, 상수값으로 표현된 C 의 값을 조정하여 support vector의 수를 조절할 수 있다. 만일 $C=1$ 인 경우 OC-SVM에서는 outlier를 허용하지 않게 되므로 모든 특징 벡터를 포함하는 최소 구를 찾게 된다. 그러므로 비디오 요약에서는 주어진 모든 특징 벡터를 다 포함하는 최소 구를 얻기 위하여 C 를 1로 가정한다. 여기서 구해진 최소 구는 구의 중심 벡터 a 와 반지름 R 로 나타내어진다. 이 때, 0보다 큰 값을 가지는 a 에 대한 특징 벡터인 서포트 벡터들은 구의 표면적에 놓이게 된다. 이렇게 표면적에 놓인 특징 벡터들의 세그먼트로부터 키 프레임을 추출하여 OC-SVM 모듈(31)은 비디오 요약에 사용한다.

<99> (4) Scalable 비디오 요약

<100> 폐지 OC-SVM의 목적이 최소 구를 찾는데 있다는 점을 이용한다. 도3의 개념도는 이를 잘 설명하고 있다.

<101> 도3을 참조하면, 우선 OC-SVM 모듈(31)이 첫 번째 표면적에 놓인 서포트 벡터를 수집하여 비디오 요약을 구성한다. 만일 이렇게 수집된 비디오 요약으로 충분하지 않다면, 도3의 개념도에서와 같이, 확장성 모듈(32)이 맨 바깥층을 구성하고 있는 구를 제거한다.(마치 양파의 겉 껍질을 벗겨내듯이) 남아 있는 특징 벡터에 다시 폐지 OC-SVM을 적용하면, 도3의 그림에서 맨 바깥쪽에서 한층 안쪽에 자리잡고 있는 구를 얻게 될 것이다.

<102> 이 새로운 구로부터 새로운 서포트 벡터(support vector)를 얻을 수 있다. 이 새로운 서포트 벡터들에 해당하는 세그먼트들을 첫 번째 구한 비디오 요약에 추가하면 조금 더 상세한

요약 정보를 얻을 수 있다. 이 때, 처음 단계의 비디오 세그먼트와 두 번째 추가된 세그먼트 간에 시작적으로 중복감을 줄 수 있으므로 이를 검증한 후 추가 시킨다. 이렇게 확장성 모듈 (32)이 가장 바깥층에 자리잡고 있는 구부터 하나씩 제거하는 작업을 원본 비디오 세그먼트가 다 소진될 때까지 혹은, 특정 조건이 만족될 때까지 계속 반복함으로써 확장 가능한 비디오 요약을 생성할 수 있다. 이러한 과정을 정리하면 다음과 같다.

<103> 입력: 중요도 정보인 β 값과 특징 벡터로 구성된 세그먼트 데이터

<104> 단계 1: 입력데이터에 대하여 식 (4')의 C=1로 하여 Quadratic Programming으로 a를 구 한다.

<105> 단계 2: 단계 1에서 $0 < a$ 의 조건을 만족하는 비디오 세그먼트의 키프레임 집합을 구 한다. 이때, 그에 해당하는 특징 벡터를 입력데이터에서 제거한다.

<106> 단계 3:

<107> Case 1 : 현재 loop에 처음 들어온 경우

<108> 단계 2에서 얻어진 키프레임들로 비디오 요약 (VS:VideoSummary)을 구성하고 단계 4로 간다.

<109> Case 2: 처음이 아니면, 세그먼트의 중요도 β 에 따라 비오름차순으로 정렬된 연속된 키프레임의 시퀀스 TS가 다 제거될 때까지 다음을 반복한다.

<110> TS에서 하나를 제거한 후, 해당 키프레임의 속한 세그먼트와 이웃하는 세그먼트가 (이웃하는 세그먼트는 분할된 원본 비디오에서 구함) 현재의 VS에 속해있는지를 확인한다.

<111> ① 속해있지 않는 경우, 방금 TS에서 제거된 keyframe을 VS에 추가한다.

<112> ② 속해있는 경우, 최소 유사도가 임계값 이하인 경우에만 추가된다.

<113> 단계 4: 단계2에서 support vector가 제거된 데이터를 입력데이터로 하여 단계 1로 간다. (혹은, 미리 고려된 임의의 특정한 종료 조건이 만족되지 않은 경우 단계 1로 돌아가 키 프레임 추출을 계속하고, 만족되는 경우 과정을 종료할 수도 있다.)

<114> 위와 같은 흐름에 따라 비디오 요약을 생성하는 Fuzzy OC-SVM시스템을 각각 영화와 뮤직비디오에 적용해 본 실험 결과를 도4 및 도5에 나타내었다.

<115> 도4와 도5의 도표에서 보여주는 것처럼, 임의의 임계값에 따라 일정하게 고정된 길이의 주요 장면들을 생성하는 것이 아니라, 몇 번의 반복 작업을 통해 동영상 비디오에서 나타나는 주요 이벤트들을 90% 이상 포함하는 주요 장면들을 구성하여 비디오 요약을 생성할 수 있다.

<116> 본 발명에서 기술된 fuzzy one-class SVM을 이용한 비디오 요약 생성 기술은 비디오 요약에 있어서 세그먼트에 대한 사용자의 주관적 중요도와 세그먼트의 시작적 특징을 동시에 고려하여 주어진 비디오의 내용을 묘사할 수 있는 최적의 세그먼트를 추출하는 기술로서, 기존의 임계치에 지나치게 의존한 비디오 요약 방식을 획기적으로 개선하여 다양한 비디오 특성에 걸 맞는 최적의 임계치를 찾아내고 이에 따라 적절한 비디오 요약을 생성한다.

<117> 이렇게 생성된 비디오 요약은 출력부(50)에 의해 화면상에 표시하고 저장부(60)는 그 출력정보를 저장한다.

<118> 한편, 본 발명은 상술한 실시 예로만 한정되는 것이 아니라 본 발명의 요지를 벗어나지 않는 범위 내에서 수정 및 변형하여 실시할 수 있다. 이상 설명한 내용을 통해 당업자라면 본 발명의 기술 사상을 일탈하지 않는 범위에서 다양한 변경과 수정 실시가 가능함을 알 수 있을 것이다.

【발명의 효과】

<119> 이상에서와 같이, 이 발명의 실시예가 응용분야에 미치는 효과는 다음과 같이 설명될 수 있다.

<120> 웹 문서 검색에 있어서 검색 엔진이 웹 페이지의 요약 정보를 제공한다. 이는 검색된 모든 결과를 일일이 방문하지 않고도 그 내용을 짐작하게 함으로써 정보를 찾는 시간을 줄이기 위한 것이다.

<121> 이와 마찬가지로, 비디오 요약의 근본적인 목적은 비디오 전체를 보지 않고도 그 내용을 파악하게 함으로써 정보 액세스의 효율을 극대화함에 있다. 비디오 정보 양을 고려할 때, 비디오 요약을 자동으로 생성하는 기술은 텍스트를 요약하는 기술보다 그 중요성이 더욱 커진다.

<122> 더욱이 대역폭이 제한되어 있는 무선 환경에서는 적은 양의 데이터로 주요 정보를 전달할 수 있는 비디오 요약 기술은 무선 비디오 컨텐츠 산업의 핵심 기술로 활용될 것이다.

【특허청구범위】

【청구항 1】

사용자로부터 요약하고자 하는 동영상 비디오 소스와 원하는 요약 시간을 입력받기 위한
입력부;

동영상 비디오의 장르별 특성 및 원하는 요약의 목적에 따라 중요도를 산출하는 중요도
측정 모듈;

샷 정보와 중요도 값을 펴지 OC-SVM(one-class support vector machines) 알고리즘에 적
용하여 비디오 요약을 생성하는 비디오 요약 생성 모듈을 포함하는 펴지 기반 OC-SVM을 이용한
동영상 자동 요약 장치.

【청구항 2】

제1항에 있어서,

전체 비디오 소스를 샷 단위로 추출하는 샷 검출 모듈을 더 포함하는 OC-SVM을 이용한
동영상 자동 요약 장치.

【청구항 3】

제1항 또는 제2항에 있어서,

생성된 비디오 요약을 화면상에 출력하는 출력부;

생성된 비디오 요약을 저장하는 저장부를 포함하는 펴지 기반 OC-SVM을 이용한 동영상
자동 요약 장치.

【청구항 4】

제3항에 있어서,

상기 비디오 요약 생성 모듈은,

샷 정보와 중요도 값을 퍼지 OC-SVM 알고리즘에 적용하여 비디오 요약을 생성하는 퍼지 OC-SVM 모듈;

사용자로부터 입력으로 받은 요약 시간 정보를 바탕으로 확장성(scalability) 처리를 반복 수행하여 사용자가 원하는 시간의 범위를 갖는 비디오 요약을 생성하는 확장성 처리 모듈을 포함하는 OC-SVM을 이용한 동영상 자동 요약 장치.

【청구항 5】

제4항에 있어서,

상기 샷 검출 모듈은 요약하고자 하는 동영상 비디오 소스로부터 샷을 검출하여 샷 리스트를 구성하고 이를 상기 비디오 요약 생성 모듈로 보내는 것을 특징으로 하는 OC-SVM을 이용한 동영상 자동 요약 장치.

【청구항 6】

사용자로부터 요약하고자 하는 동영상 비디오 소스와 원하는 요약 시간을 입력받는 제1 단계;

상기 비디오 소스를 샷 단위로 추출하는 제2 단계;

동영상 비디오의 장르별 특성 및 원하는 요약의 목적에 따라 중요도를 산출하는 제3 단계;

샷 정보와 중요도 값을 퍼지 OC-SVM 알고리즘에 적용하여 비디오 요약을 생성하는 제4 단계를 포함하는 OC-SVM을 이용한 동영상 자동 요약 방법.

【청구항 7】

제6항에 있어서,

생성된 비디오 요약을 화면상에 출력하는 제5 단계;

생성된 비디오 요약을 저장하는 제6 단계를 더 포함하는 OC-SVM을 이용한 동영상 자동 요약 방법.

【청구항 8】

제6항 또는 제7항에 있어서,

상기 제4단계는 샷 정보와 중요도 값을 펴지 OC-SVM 알고리즘에 적용하여 비디오 요약을 생성하고, 사용자로부터 입력으로 받은 요약 시간 정보를 바탕으로 확장성(scalability) 처리를 반복 수행하여 사용자가 원하는 시간의 범위를 갖는 비디오 요약을 생성하는 것을 특징으로 하는 OC-SVM을 이용한 동영상 자동 요약 방법.

【청구항 9】

사용자로부터 요약하고자 하는 동영상 비디오 소스와 원하는 요약 시간을 입력받는 제1 단계;

동영상 비디오의 장르별 특성 및 원하는 요약의 목적에 따라 중요도를 산출하는 제2 단계;

샷 정보와 중요도 값을 펴지 OC-SVM 알고리즘에 적용하여 비디오 요약을 생성하는 제3 단계;

생성된 비디오 요약을 화면상에 출력하는 제4 단계;

생성된 비디오 요약을 저장하는 제5 단계를 더 포함하는 OC-SVM을 이용한 동영상 자동 요약 방법.

【청구항 10】

OC-SVM을 이용한 동영상 자동 요약 방법을 기록한 기록매체로서,
사용자로부터 요약하고자 하는 동영상 비디오 소스와 원하는 요약 시간을 입력받는 제1 단계;

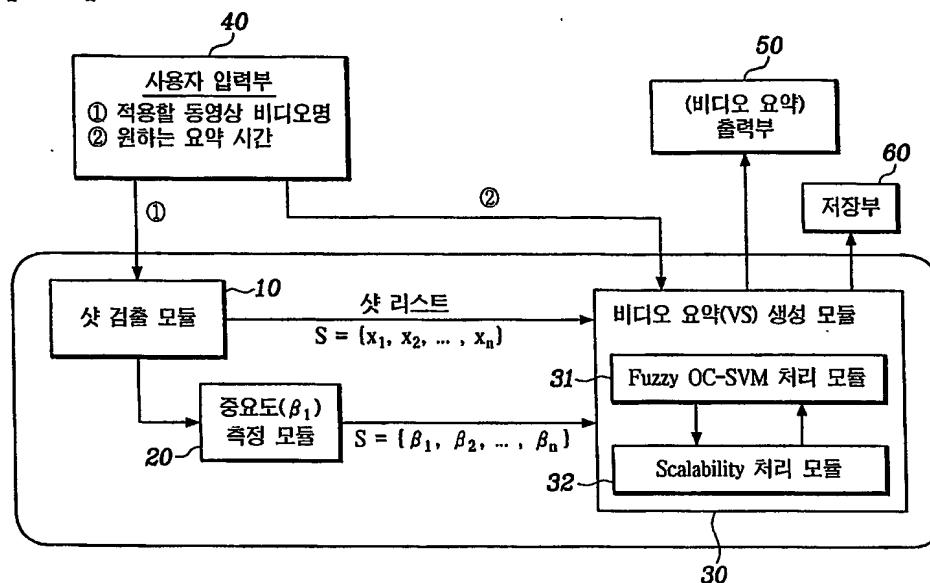
상기 비디오 소스를 샷 단위로 추출하는 제2 단계;

동영상 비디오의 장르별 특성 및 원하는 요약의 목적에 따라 중요도를 산출하는 제3 단계;

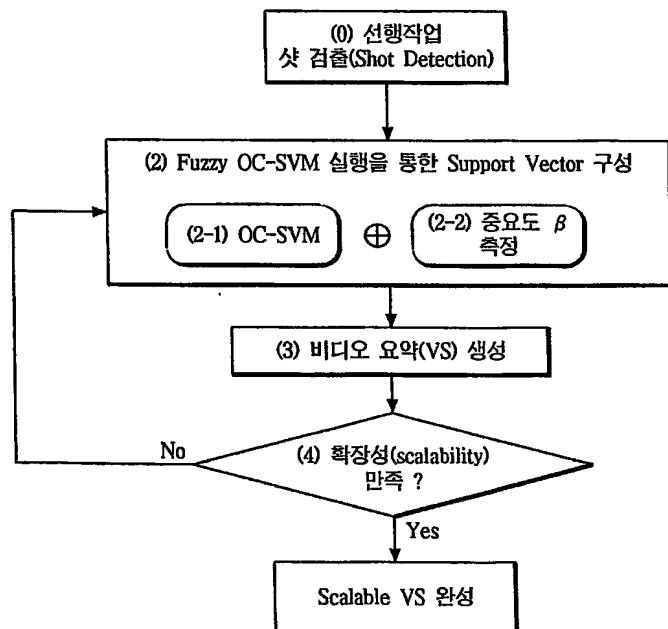
샷 정보와 중요도 값을 퍼지 OC-SVM 알고리즘에 적용하여 비디오 요약을 생성하는 제4 단계를 포함하는 기능이 구현된 프로그램을 저장한 기록매체.

【도면】

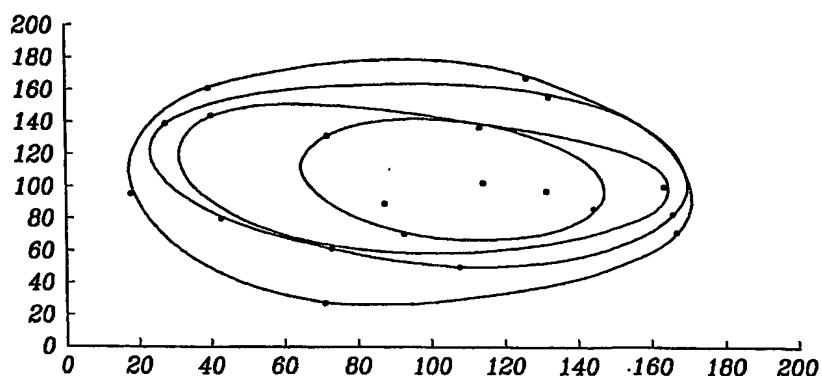
【도 1】



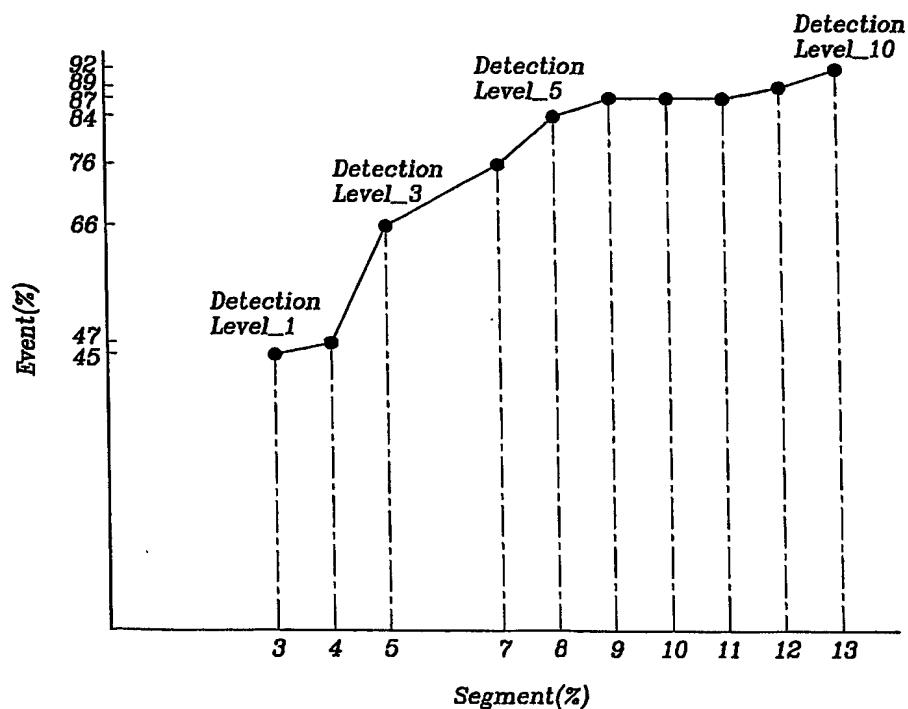
【도 2】



【도 3】



【도 4】



1020 142613

출력 일자: 2004/7/5

【도 5】

